ATTORNEY DOCKET NO.: 10010310-1

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

This is a U.S. Patent Application for:

Title: **MEANING TOKEN DICTIONARY FOR AUTOMATIC SPEECH RECOGNITION**

Inventor: MICHAEL VANHILST
Address: PO Box 800616, Aventura, FL 33280-0616
Citizenship: United States

MEANING TOKEN DICTIONARY FOR AUTOMATIC SPEECH RECOGNITION

TECHNICAL FIELD

This invention relates to a meaning token dictionary for automatic speech recognition applications, including task-oriented speech recognition applications.

BACKGROUND

Most real-time automatic speech recognition systems include a speech recognizer and a natural language processor. The speech recognizer converts an incoming stream of sounds into a likely sequence of words. In particular, the speech recognizer receives digitized speech samples from an acoustic input device (e.g., a microphone), and converts the digitized speech samples into sequences of recognized words based upon finite state machine templates. The finite state machine templates are defined by a set of vocabulary word patterns, which are stored in a dictionary, and, possibly, a set of grammar rules. The natural language processor attempts to make sense of the word sequences that are output by the speech recognizer. That is, the natural language processor attempts to extract from each word sequence a meaning that is relevant to a specific user application. In a typical implementation, the natural language processor compares a word sequence with internal semantic patterns that are defined by a grammar compiler. The natural language processor identifies permissible sentences based upon the internal semantic patterns, and outputs summary structures representing permissible sentences. In task-oriented automatic speech recognition systems, an application command translator may match the summary structures with a known set of user application commands, and may return appropriate commands to a user application for processing.

Speech-enabled applications that understand normal spoken sentences often are difficult to implement and usually exhibit frequent recognition errors. In general, the process of automatically recognizing speech is made difficult by three primary obstacles. First, most words are short. With only a few sound features, a speech recognizer usually has difficulty in clearly distinguishing among similar sounding

candidates. As a result, numerous incorrect hypotheses from the dictionary are likely to be passed on to the grammar process. Second, grammar processes that attempt to make fine distinctions among possible utterances are complicated, inefficient, and hard to write. Simpler grammars, on the other hand, allow many incorrect utterances that cannot be parsed (or interpreted) for meaning. Finally, even if a correct sentence is transcribed and output, natural language itself typically is ambiguous. Thus, the parsing process implemented by the user application often must test numerous hypotheses using complicated, non-deterministic methods to determine the correct meaning from a natural language input.

SUMMARY

The invention features systems and methods of automatic speech recognition in which the vocabulary words of a speech recognition dictionary correspond to meaning tokens each of which signifies a single meaning.

In one aspect, the invention features an automatic speech recognition system that comprises a speech recognition dictionary and a speech recognizer. The speech recognition dictionary comprises a plurality of meaning tokens each associated with one or more pronunciations of one or more vocabulary words and signifying a single meaning. The speech recognizer is configured to convert spoken input into a sequence of meaning tokens that are contained in the speech recognition dictionary and correspond to a sequence of vocabulary words that are most likely to have been spoken by a user.

Embodiments of the invention may include one or more of the following features.

Each meaning token preferably is characterized by a unique spelling. The spelling of a meaning token preferably facilitates extraction of meaning by a language analyzer. The spelling of a meaning token may encode one or more labels identifying one or more respective application-specific categories. An application-specific category that is identified by a label encoded in the spelling of a meaning token may be an object category, a place category, an event category, or an action category.